OC3140
HW/Lab 7 Hypothesis Testing

1. Is the sea surface temperature in September in the Gulf of Mexico significantly hotter than $25^o C$? A sample of 10 days of SST was taken (sample size n = 10). The sample mean and standard deviation are $27^o C$ and $2^o C$, respectively.

Solution:

This is the Hypothesis Testing on the mean (see Ch. 7 p7-p8) (as small sample and unknown s)

- Start on the null hypothesis: H_0 : $m_{sst} = 25^o C$
- Define the alternative hypothesis: H_A : $m_{sst} > 25^o C$
- Identify a test statistic: for a testing on means, a t distribution is used:
$$t = \frac{\overline{x} - m_{sst}}{s/\sqrt{n}}$$
- Identifying a critical value t_c with $a = 0.05$, $d.f. = n - 1 = 9$ and H_A is upper one-side, we have (from the t table Ch.5 p17)
$$t_c = 1.833.$$
- For the testing sample,
n=10, mean=*27*, standard deviation = *2*,
the *t*-value is
$$t = \frac{27-25}{2/\sqrt{10}} = 3.1623.$$
- Since $t(=3.1623) > t_c(=1.833)$, we reject the null hypothesis H_0 and conclude that sea surface temperature is significantly hotter than $25^o C$.

2. A marine science equipment needs dozens uniform batteries. The manufacturer claims a variance of 0.012. Sample size = 50 batteries, sample variance=0.02. Determine if the manufacture's claim can be accepted.

Solution:

This is the Hypothesis Testing on the Variance (see Ch. 7, p8-p9)

- Start on the null hypothesis: H_0: $s^2 = 0.012$
- Define the alternative hypothesis: H_A: $s^2 > 0.012$
- Identify the test statistics,

$$c^2 = \frac{(n-1)\,s^2}{s^2}$$

- Identifying a critical value base on:
  $a = 0.05,\ d.f. = n-1 = 49$ and $H_a:\ s^2 > 0.012$ (i.e., upper one-sided),
  we have (from the c^2 table Ch.5 p12),
  $$c_c^2 = c_{0.05,49}^2 = 67.5.$$
- Compute the test statistics from the sample:
  $$c^2 = \frac{(n-1)\,s^2}{s^2} = \frac{49 \cdot 0.02}{0.012} = 81.67.$$
- Since $c^2\,(=81.67) > c_c^2\,(=67.5)$, we reject H_0 and conclude that the batteries are significantly more variable than the manufacture claimed.

3. The quality control requires a product from a factory to have no more than 2% defective rate. When 300 units are randomly taken from the factory, it finds that 5 of them are defective. Can the produce meet the quality control requirement? A significance level of 5% is used.

Solution:

This is the Hypothesis Testing on the Proportion (see Ch. 7 p9-p10)

- State the null hypothesis: $H_o:\ p{=}0.02$
- Define the alternative hypothesis: $H_A:\ \ p \le 0.02$
- Identify the test statistics: for a testing on mean, a z distribution is used.
  $$z = \frac{n \cdot p - n \cdot p_o}{\sqrt{n \cdot p_o\,(1-p_o)}}$$
- Identifying a critical value for z, z_c: Based on (a) $a = 0.05$ and (b) H_A is high one-sided, we have $z_c = 1.645$ (Ch III page 23)
- Compute the test statistic from the sample:
  $$z = \frac{300 \cdot \left(\frac{5}{300}\right) - 300 \cdot 0.02}{\sqrt{300 \cdot 0.02 \cdot (1-0.02)}} = -0.4124$$
- Since $z\,(=-0.4124) < z_c\,(=1.645)$, the z value is not in the critical region. Thus we accept the null hypothesis H_o and accept the alternative hypothesis: H_A. In other words, the product from the factory meets the QC requirement.

4. There are some temperature samples of this month in Monterey city and Marina city. The sample size in Monterey is 25 and the ample size in Marina city is 17. The sample means and variances were computed and tabulated in the following.

	Sample Size	Sample mean	Sample Variance
Monterey	25	75.5	4
Marina	17	73.3	6

Is there a significant difference of the mean temperature between Monterey and Marina cities at a significance level of 5%?

Solution:

This is the Hypothesis Testing on the difference of two means (see Ch. 7 p10-p12)

- State the null hypothesis: $H_o:\ m_1 = m_2$
- Define the alternative hypothesis: $H_A:\ m_1 \neq m_2$
- Identify the test statistics: t distribution.

Method: 1 assume $s_1 = s_2$

- Identifying a critical value: $d.f. = n_1 + n_2 - 2 = 40$, $a = 0.05$ and it is two-sided testing, we have $t_{\frac{a}{2},40} = 2.0211$
- Compute the t value from the sample

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{75.5 - 73.3}{\sqrt{\left(\frac{24*4 + 16*6}{40}\right)\cdot\left(\frac{1}{25} + \frac{1}{17}\right)}} = 3.1943 \text{ where}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- The t value from sample (3.1943) is greater than the critical value (2.0211). We reject the null hypothesis (accept the alternative hypothesis) in other words, this month Monterey and Marina have different mean temperature.

Method: 2 assume $s_1 \neq s_2$

- Identifying a critical value with a degree-of-freedom of

$$\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2 = \frac{\left(\frac{4}{25} + \frac{6}{17}\right)^2}{\left(\frac{\left(\frac{4}{25}\right)^2}{26} + \frac{\left(\frac{6}{17}\right)^2}{18}\right)} - 2 = 31.2837 \approx 31$$

$$t_{0.025,31} = \frac{1}{10}\left(9 * t_{0.025,30} + 1 * t_{0.025,40}\right) = 2.0399$$

- Compute the t value from the sample

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{75.5 - 73.3}{\sqrt{4/25 - 6/17}} = 3.0718$$

- The t value from sample (3.0718) is greater than the critical value (2.0399).

We reject the null hypothesis (accept the alternative hypothesis) in other words, this month Monterey and Marina have different mean temperature